

Chapitre I

Introduction aux spectroscopies optiques et à la diffusion des neutrons et des rayons X.

La diffusion de neutrons et de rayons X sont les principales techniques qui ont été employées au cours de ce travail sur les états dénaturés de la NCS. Dans ce chapitre, nous donnons quelques principes nécessaires à la compréhension de nos études. Nous exposons tout d'abord les spécificités de la diffusion des neutrons et des rayons X aux petits angles et les lois utilisées pour analyser les résultats. Ensuite, nous présentons les points les plus importants de la diffusion incohérente inélastique des neutrons, les détails concernant l'analyse de nos données étant exposés dans le chapitre 5. Nous rappelons aussi très rapidement les concepts de base des autres techniques utilisées comme la fluorescence, le dichroïsme circulaire, la diffusion de lumière et la microcalorimétrie.

1.1 Les spectroscopies optiques

1.1.1 La fluorescence.

En absorbant un photon, un fluorophore passe de son état électronique fondamental à un état excité, instable. Le système retourne à l'état fondamental en **émettant un photon**. En général, l'émission se produit très rapidement après l'excitation. Sa durée de vie est de l'ordre de la nanoseconde (0.1-100 ns) (phénomène de fluorescence). La grandeur mesurée pendant une expérience de fluorescence est l'intensité d'émission (ou de fluorescence), d'un système qui a été excité, en fonction de la longueur d'onde de la lumière réémise.

Le signal de fluorescence dépend des interactions directes et indirectes du fluorophore avec son environnement. Avec cette technique, il est possible d'étudier les changements de conformation des protéines en solution, c'est-à-dire dans des conditions similaires aux conditions physiologiques. Les protéines fluorescent naturellement dans l'ultraviolet (UV) proche. Parmi les vingt acides aminés qui les composent, seuls trois résidus possédant des structures aromatiques, peuvent réémettre la lumière (phénylalanine, tyrosine et tryptophane). La fluorescence du tryptophane est, de beaucoup, la plus intense et est très sensible à son environnement. En général, les protéines ne contiennent que très peu de tryptophanes mais leur contribution domine le spectre de fluorescence. Une caractéristique du signal de fluorescence est donc dépendre essentiellement de quelques groupes fluorescents, qui ont une taille petite par rapport aux dimensions de la protéine. Cela implique que les informations que nous allons obtenir sont des informations locales qui concernent l'environnement de chaque fluorophore.

La lenteur relative du phénomène d'émission permet au fluorophore d'échanger plus ou moins d'énergie avec son environnement. Lorsqu'une protéine change de conformation chaque fluorophore se trouve un environnement différent et fluoresce donc de manière différente. La somme des contributions des différents fluorophores atténue la structure vibrationnelle du spectre de fluorescence, et le rend continu. En général, lorsque les tryptophanes sont enfouis dans la matrice protéique ils donnent un spectre plus structuré que

lorsqu'ils sont exposés au solvant. De plus ce spectre est décalé vers les courtes longueurs d'ondes.

Il est aussi possible d'utiliser une sonde fluorescente extérieure à la protéine. Certaines molécules, comme les sulfonates d'anilino-naftalène (ANS), ne fluorescent que dans un environnement apolaire. En ajoutant ce composé à une solution de protéine et en étudiant sa fluorescence il devient donc possible de déterminer dans quelles conditions une protéine se déstructure suffisamment pour devenir perméable au solvant. Cette méthode est souvent employée pour mettre en évidence des états de type *molten globule* et plus généralement des états dénaturés compacts qui sont, en général, très difficiles à détecter par d'autres méthodes.

1.1.2 Dichroïsme circulaire

Les molécules biologiques sont des molécules *chirales*, c'est-à-dire sans centre ni plan de symétrie. Elles présentent donc une activité optique. Le dichroïsme circulaire est un aspect de cette activité. Si on considère l'interaction de la lumière avec une solution de molécules biologiques, on peut observer que les vitesses de propagation de la lumière polarisée circulairement droite et gauche sont différentes. Il s'ensuit que la lumière, après avoir traversé la solution de protéine, est polarisée elliptiquement. L'ellipticité peut être positive ou négative. Il est d'usage de l'exprimer sous la forme **d'ellipticité molaire** : $\theta = 2.303 \cdot \Delta A / 4l$, où ΔA , qui représente l'observable mesurée, est la différence d'absorbance entre les polarisations droite et gauche, et l est le trajet optique.

Le dichroïsme circulaire des bandes d'absorption des liaisons peptidiques s'observe principalement entre 200 et 300 nm. Il permet de distinguer les principaux types de structure, car chaque type de structure secondaire a un spectre caractéristique (Fig 1.1)

Dans la région de l'ultraviolet lointain, au dessous de 250 nm, les spectres dépendent essentiellement de la conformation du squelette peptique et, en particulier, des différents types de structure secondaire. Le dichroïsme s'annule presque complètement quand la protéine est totalement dénaturée. Les protéines natives ont aussi une activité optique significative dans le proche ultraviolet, entre 250 nm et 300 nm. Elle est essentiellement due aux groupes aromatiques situés dans des environnements asymétriques et à la chiralité induite par les liaisons disulfures. A une longueur d'onde fixe, la dépendance de l'ellipticité molaire d'une

protéine, en fonction par exemple de la température ou du pH, donne des d'informations sur l'évolution des structures secondaires de la protéine au cours de la dénaturation.

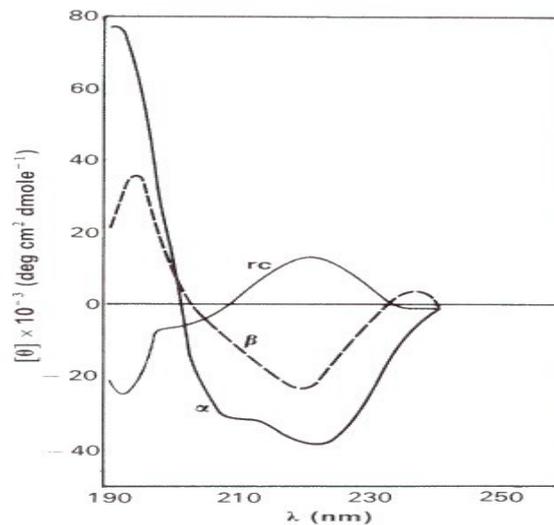


Fig 1.1 Spectre de dichroïsme circulaire caractéristique pour une hélice α , un feuillet β , et un random coil (Creighton 1993).

1.1.3 Analyse thermodynamique des résultats

La variation de ces quantités, au cours de la transition de dénaturation d'une protéine, permet d'extraire non seulement des informations structurales, mais aussi certains paramètres thermodynamiques. En particulier, dans le cas des processus de dépliement qui peuvent être décrit en première approximation par un modèle à deux états (présence en solution de deux états distincts, l'état natif (N) et l'état dénaturé (D)) nous pouvons évaluer la variation d'énergie libre, ΔG^0 , et l'enthalpie, ΔH , associée à ces processus. Désignons par K_D la constante d'équilibre de la transition $N \rightarrow D$, par f_D la fraction de protéine dénaturée, par ΔH^0 et ΔS^0 les variations d'enthalpie et d'entropie associées à la dénaturation. Ces quantités sont liées par les relations suivantes:

$$K_D = \frac{f_D}{1 - f_D} \quad (1.1)$$

et

$$\Delta G^0 = -RT \ln K_D = \Delta H^0 - T\Delta S^0 \quad (1.2)$$

où R est la constante des gaz parfaits et T la température.

Lors d'une dénaturation thermique, ces grandeurs thermodynamiques dépendent de la température [Privalov 1979, 1989]. Selon la relation de Gibbs- Helmholtz la variation d'énergie libre à la température T , s'écrit :

$$\Delta G^0(T) = \Delta H(T_0) + \Delta C_P(T - T_0) - T[\Delta S(T_0) + \Delta C_P \ln(T / T_0)] \quad (1.3)$$

où ΔC_P est la variation de chaleur spécifique apparaissant lors de la dénaturation. $\Delta H(T_0)$ et $\Delta S(T_0)$ sont les valeurs des variations d'enthalpie et d'entropie à la température de référence T_0 , qui peut être la température de demi-transition, c'est-à-dire celle à laquelle les populations des deux états sont égales.

Pour une dénaturation induite par un agent chimique, comme le chlorure de guanidinium (GdmCl), la variation d'énergie libre dépend de la concentration, $[D]$, de dénaturant. En première approximation cette dépendance peut être considérée comme linéaire [Tanford 1968, Schellman 1978, Pace 1986] :

$$\Delta G^0([D]) = \Delta G^0 - m[D] \quad (1.4)$$

où $[D]$ est la molarité du dénaturant et m est une constante proportionnelle à la surface de la protéine devenue accessible au solvant lors de la dénaturation.

A partir de ces considérations nous pouvons définir une expression décrivant la variation du signal expérimental en fonction de la température et de la concentration de dénaturant. La variation du signal expérimental, $Y(X)$ en fonction d'un paramètre X , peut s'écrire :

$$y(X) = [1 - f_D(X)]y_N(X) + f_D(X)y_D(X) \quad (1.5)$$

Pour tenir compte des effets du solvant, nous supposons que le signal correspondant à l'état natif, $Y_N(X)$, et à l'état dénaturé, $Y_D(X)$, dépend linéairement de X . En utilisant la définition (1.1) l'expression précédente peut donc s'écrire :

$$y(X) = y_N + O_N X + \frac{K_D(X)}{1 + K_D(X)} [(y_N - y_D) + (O_N - O_D)X] \quad (1.6)$$

où O_N et O_D sont des constantes D'après la définition (1.2)

$$\frac{K_D(X)}{1 + K_D(X)} = \frac{\exp[\Delta H(T_0)(1/RT_0 - 1/RT)]}{1 + \exp[\Delta H(T_0)(1/RT_0 - 1/RT)]} \quad (1.7)$$

pour une dénaturation thermique et dans le cas d'une dénaturation chimique :

$$\frac{K_D(X)}{1 + K_D(X)} = \frac{\exp(\Delta G^0 - m[D])/RT}{1 + \exp(\Delta G^0 - m[D])/RT} \quad (1.8)$$

Les paramètres thermodynamiques $\Delta H(T_0)$, d'une part, et ΔG^0 et m , d'autre part, sont obtenus à partir de l'ajustement de ces trois dernières équations aux données expérimentales. Cette méthode a été utilisée pour analyser les courbes de transition de fluorescence et de dichroïsme circulaire. Elle a aussi été employée pour décrire les variations du carré du rayon de giration de la NCS en fonction de la température et de la concentration en GdmCl.

1.2 La microcalorimétrie différentielle par balayage

1.2.1 Principe de la technique

L'importance des mesures de calorimétrie pour l'étude du dépliement d'une protéine est principalement liée au fait que c'est la seule méthode capable de déterminer directement l'enthalpie et la variation de la chaleur spécifique associées au processus d'intérêt. Au moyen une modélisation, ces paramètres peuvent être mis en relation avec les modifications structurales et les mécanismes sous-jacents. (Freire, 1994, 1995).

La calorimétrie différentielle par balayage permet de mesurer la capacité calorifique d'un ensemble de macromolécules en fonction de la température. Si ces molécules subissent une ou plusieurs transitions structurales induites par la température, la capacité calorifique moyenne, $\langle C_p(T) \rangle$ présentera un ou plusieurs pic centrés à des températures voisines des températures de demi-transition T_M . Il est d'usage de considérer que $\langle C_p(T) \rangle$ est la somme de deux contributions :

$$\langle C_p(T) \rangle = \langle \Delta C_{p, tr}(T) \rangle + \langle \Delta C_{p, bl}(T) \rangle \quad (1.9)$$

où $\langle \Delta C_{p, tr}(T) \rangle$ est l'excès de capacité calorifique correspondant à la chaleur latente de transition et $\langle \Delta C_{p, bl}(T) \rangle$ représente la variation de capacité calorifique qu'aurait le système en l'absence de chaleur latente [Freire 1994, 1995]. La différence entre la chaleur spécifique de l'état natif et de l'état dénaturé de la protéine est indiquée par la fonction $\Delta C_p(T)$. La Fig 1. 2 présente une courbe typique de $C_p(T)$.

En général, la variation moyenne $\langle \Delta C_p(T) \rangle$ de la chaleur spécifique est obtenue en soustrayant la chaleur spécifique de l'état natif, $C_{p, n}(T)$, de la chaleur spécifique mesurée $\langle C_p(T) \rangle$ [Freire1994, 1995]:

$$\langle \Delta C_p(T) \rangle = \langle C_p(T) \rangle - C_{p, n}(T) \quad (1.10)$$

on suppose $C_{p,n}(T)$ linéaire en fonction de la température. Dans le chapitre II, nous verrons de façon plus détaillée comment les résultats bruts sont traités.

A partir de l'excès de capacité calorifique $\langle \Delta C_{p,tr}(T) \rangle$, il est possible de calculer la variation d'enthalpie correspondante, $\langle \Delta H \rangle$. Elle est égale à la surface délimitée par le pic de transition, et correspond à la somme des contributions relatives de tous les états qui sont peuplés pendant la transition:

$$\langle \Delta H \rangle = \int_{T_0}^{T_f} \langle \Delta C_{p,tr} \rangle dT \quad (1.11)$$

où T_0 et T_f sont les températures initiale et finale, respectivement. La variation d'entropie correspondante est :

$$\langle \Delta S \rangle = \int_{T_0}^{T_f} \langle \Delta C_{p,tr} \rangle d \ln T \quad (1.12)$$

La variation d'enthalpie associée au processus de dénaturation est due à plusieurs facteurs, comme la rupture des interactions intramoléculaires (interactions de van des Waals, liaisons hydrogène,...) et la formation concomitante de nouvelles interactions avec le solvant (protonation, ionisation,...). Par contre, la variation d'entropie provient essentiellement de la perte de structure de la protéine et du transfert des chaînes latérales de l'intérieur de la protéine vers le solvant (Freire 1994, 1995).

Enfin, la chaleur spécifique absolue d'une protéine complètement dénaturée $C_{p,u}(T)$ peut être calculée à partir de la séquence primaire, tous les résidus étant exposés, $C_{p,u}(T)$ suit simplement la règle de l'additivité et peut être écrite comme la somme des contributions individuelles de tous les résidus latéraux et du squelette. Privalov et Makhatadze (1990) ont calculé les paramètres nécessaires pour estimer la valeur de $C_{p,u}(T)$ à partir de la séquence des acides aminés, modélisant une dépendance cubique de $C_{p,u}(T)$ avec la température.

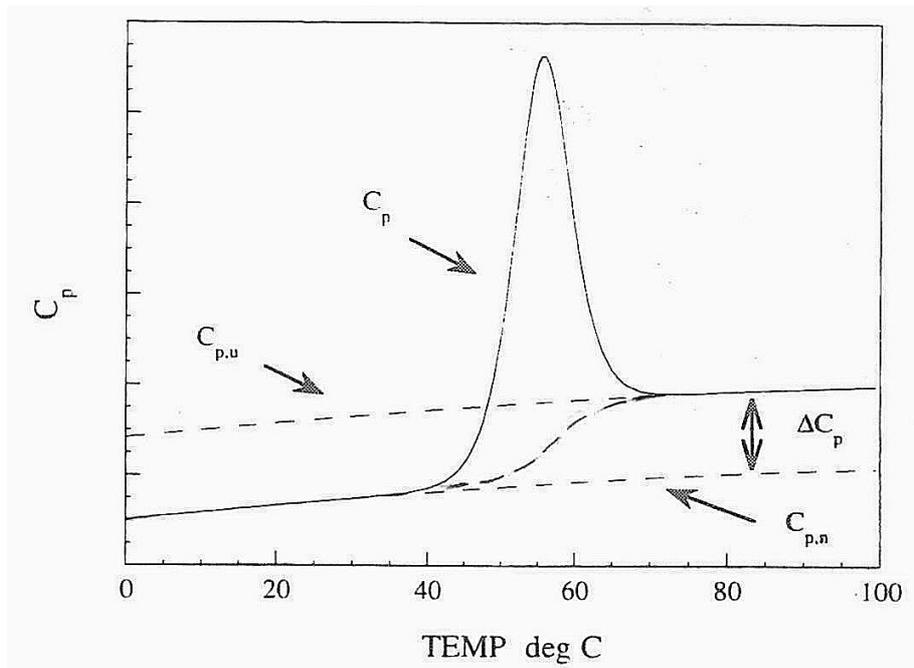


Fig 1. 2 Variation de la chaleur spécifique molaire C_p , représentant la transition entre l'état natif et l'état dénaturé d'une protéine. $C_{p,u}$ et $C_{p,n}$ indiquent la chaleur spécifique molaire de l'état dénaturé et de l'état natif respectivement. ΔC_p la différence de chaleur spécifique entre l'état natif et dénaturé.

La transition sigmoïdale, entre $C_{p,u}$ et $C_{p,n}$ représente $\Delta C_{p,bl}$. Le pic représente $\Delta C_{p,ir}$

1.2.2 Enthalpie de van't Hoff

L'enthalpie calorimétrique, ou molaire, c'est-à-dire celle qu'on mesure directement, représente la quantité de chaleur réellement absorbée lors de la dénaturation (eq. 1.11), quel que soit le nombre d'états intermédiaires impliqués dans la transition. Par contre, l'enthalpie de Van't Hoff, $\langle \Delta H_v \rangle$, est l'enthalpie associée à la transition, tenant compte seulement de l'état initial et final. On suppose que seuls deux états différents coexistent lors de la dénaturation. L'enthalpie de Van't Hoff est définie par :

$$\ln K_D = (-\Delta H_v / RT) + \Delta S / R \quad (1.13)$$

Pratiquement, l'enthalpie de Van't Hoff est calculée à partir des mesures en utilisant la relation :

$$\langle \Delta H_v \rangle = \frac{4RT_M^2}{\langle \Delta H \rangle} \langle \Delta C_p(T_M) \rangle \quad (1.14)$$

où $\langle \Delta C_P(T_M) \rangle$ est la valeur maximale de la chaleur spécifique atteinte à la température de demi-transition T_M .

Habituellement, l'évaluation du caractère à deux états d'une transition est faite en comparant les valeurs des deux enthalpies. Si elles sont égales la transition peut être assimilée à une transition à deux états. Dans le cas contraire, la présence d'un ou plusieurs états intermédiaires doit être envisagée. Une enthalpie de Van't Hoff inférieure à l'enthalpie calorimétrique indique que les états intermédiaires sont d'origine intramoléculaire. C'est ce qui se produit lorsqu'une protéine possède deux ou plusieurs domaines de stabilité différente. Par contre, une enthalpie calorimétrique plus petite que celle de Van't Hoff traduit la présence d'effets intermoléculaires tels que l'oligomérisation ou l'agrégation [Privalov 1979, Freire 1994].

En général, ce type d'analyse n'est pas très précis car la détermination de la valeur de l'enthalpie de Van't Hoff n'utilise qu'un seul point du thermogramme. De plus la présence d'états intermédiaires ayant une enthalpie proche soit de celle de l'état natif, ou de celle de l'état dénaturé, ne modifierait pas de manière significative le rapport des deux enthalpies. Une analyse plus complète des thermogrammes est donc nécessaire pour déterminer si la dénaturation implique seulement deux états ou plus.

1.3 Les techniques de diffusion

Les techniques de diffusion du rayonnement permettent d'obtenir d'importants renseignements structuraux sur les diffuseurs. Lorsque ceux-ci sont flexibles, comme les protéines fortement dénaturées, ce sont les seules, avec la R.M.N. qui puissent donner de telles informations. Depuis quelque temps la diffusion des rayons et des neutrons commencent à être utilisées pour caractériser les différentes conformations des protéines dénaturées. Dans ce paragraphe nous expliquerons d'abord le principe d'une expérience de diffusion. Nous donnerons ensuite quelques éléments de la théorie de la diffusion élastique et quasi-élastique des neutrons. Enfin, nous traiterons brièvement le cas des rayons X.

1.3.1. Principe d'une expérience de diffusion

Comme illustré dans la Fig 1. 3, le principe d'une expérience de diffusion consiste à envoyer sur l'échantillon un rayonnement monochromatique et à mesurer la répartition spatiale de l'intensité diffusée par l'échantillon. La direction incidente du rayonnement est spécifiée par le vecteur d'onde \mathbf{k}_i de module $k_i = 2\pi/\lambda$ où λ est la longueur d'onde de rayonnement dans l'échantillon. Le rayonnement diffusé est détecté dans la direction définie par le vecteur de diffusion \mathbf{k}_f faisant un angle θ avec \mathbf{k}_i .

Le **transfert de vecteur d'onde** correspondant est :

$$\mathbf{Q} = (\mathbf{k}_i - \mathbf{k}_f).$$

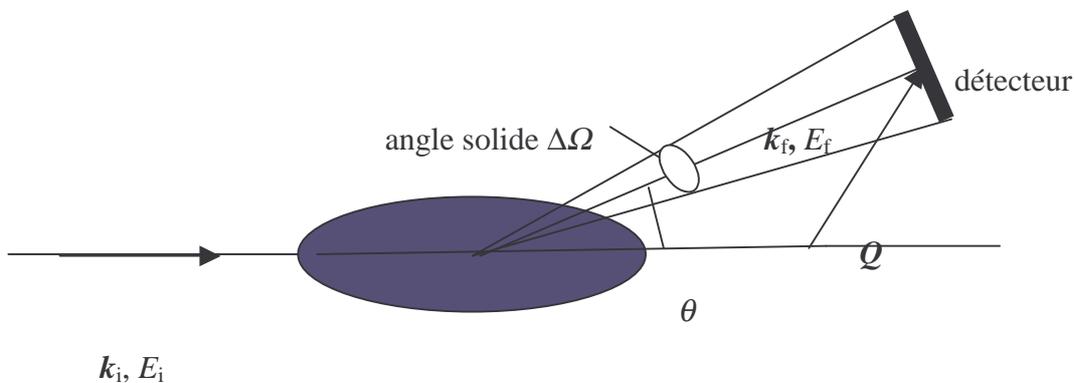


Fig 1. 3 Schéma d'une expérience de diffusion.

Lors de la diffusion, le rayonnement échange aussi de l'énergie avec le milieu. Le transfert d'énergie est $\Delta E = (E_i - E_f)$ où E_i et E_f sont, respectivement, les énergies du rayonnement incident et diffusé.

Si la diffusion est élastique, l'énergie des rayonnements incident et diffusé sont identiques. Alors $k_i = k_f$ et le transfert de vecteur d'onde :

$$Q = 2 k_i \sin(\theta/2).$$

De façon tout à fait générale, le spectre d'énergie de l'intensité diffusée par unité d'angle solide, Ω , et pour un transfert en énergie $\omega \pm d\omega$ s'écrit :

$$I_f(\mathbf{Q}, \omega) = I_0 \frac{d^2\sigma(\mathbf{Q}, \omega)}{d\Omega d\omega} \quad (1.15)$$

où I_0 est l'intensité incidente et $d^2\sigma(\mathbf{Q}, \omega)/d\Omega d\omega$ la **section efficace double différentielle de diffusion** du système. $I_f(\mathbf{Q}, \omega)$ est une quantité qui reflète à la fois la structure et la dynamique d'un système. Lorsque seule la structure du diffuseur importe, l'analyse de l'intensité diffusée en énergie n'est plus nécessaire et l'on mesure seulement l'intensité diffusée par unité d'angle solide

$$I_f(\mathbf{Q}) = I_0 \frac{d\sigma(\mathbf{Q})}{d\Omega} \quad (1.16)$$

où

$$\frac{d\sigma(\mathbf{Q})}{d\Omega} = \int \frac{d^2\sigma(\mathbf{Q}, \omega)}{d\Omega d\omega} d\omega. \quad (1.17)$$

La **section efficace totale** de diffusion :

$$\sigma = \int \frac{d\sigma(\mathbf{Q})}{d\Omega} d\Omega. \quad (1.18)$$

représente le rapport de l'intensité totale diffusée à l'intensité incidente.

Les différents types d'interaction des rayonnements (neutron, rayons X, lumière) avec la matière caractérisent les expressions des sections efficaces.

1.3.2 Diffusion des neutrons

Le neutron est une particule subatomique de charge nulle et masse $1.66 \cdot 10^{-24}$ g, de spin $\frac{1}{2}$ et moment magnétique $\mu = 1.9132 \mu_N$. La longueur d'onde associée à un faisceau de neutrons est définie par la relation de de Broglie et son énergie correspond à l'énergie cinétique :

$$\lambda = \frac{h}{mv}$$

$$E = \frac{1}{2}mv^2 = \frac{h^2}{2m\lambda^2} = \frac{\hbar^2 k^2}{2m}$$

(1. 19)

Les longueurs d'onde associées aux neutrons froids et thermiques¹ sont de l'ordre de grandeur des distances interatomiques tandis que leurs énergies sont comparables à celles des mouvements intramoléculaires. Ces propriétés font des neutrons une sonde particulièrement adaptée à l'étude de la matière condensée.

Le neutron interagit principalement avec les noyaux des atomes. La force d'interaction entre un neutron et un noyau dépend de l'élément considéré. Elle est caractérisée par une grandeur appelée **longueur de diffusion**, dénommée b . Le potentiel d'interaction nucléaire est un potentiel central, qui dans le cadre de l'approximation de Born est décrit par le pseudo-potential de Fermi (Squires, 1978). On peut dire qu'il y a interaction lorsque le neutron pénètre dans la sphère de rayon b centrée sur l'atome. En dehors de cette sphère il n'y a pas d'interaction.

La section efficace de diffusion traduisant l'interaction des neutrons avec un atome isolé de longueur de diffusion b sera égale à

¹ Les neutrons thermiques possèdent une énergie de l'ordre de grandeur de l'agitation thermique à 300K soit environ 25meV.

$$\frac{d\sigma}{d\Omega} = b^2 \quad (1.20)$$

Lors de la diffusion, le neutron échange avec le système l'énergie :

$$\hbar\Delta\omega = \hbar(\omega_i - \omega_f) = \frac{\hbar^2}{2m}(k_i^2 - k_f^2) \quad (1.21)$$

et le transfert de moment associé est :

$$\begin{aligned} Q^2 &= k_i^2 + k_f^2 - 2k_i k_f \cos\theta \\ &= k_i^2 \left(2 + \frac{\omega_f}{\omega_i} \right) - 2k_i^2 \cos\theta \sqrt{\frac{\omega_f}{\omega_i} + 1} \end{aligned} \quad (1.22)$$

Lorsque la diffusion peut être considérée comme élastique, $\omega_i = \omega_f$ et le transfert de moment est :

$$Q = \frac{4\pi}{\lambda} \sin\theta / 2 \quad (1.23)$$

Cette simplification est applicable dans le cas de la diffusion aux petits angles et, en première approximation, dans le cas de la diffusion quasi-élastique.

Nous allons développer plus en détail dans les prochains paragraphes les principes de base qui régissent l'interaction neutron matière, et discuter de la théorie de la diffusion.

1.3.3.1 Diffusion cohérente et incohérente.

La longueur de diffusion b est une grandeur indépendante de l'énergie du neutron mais qui varie en fonction du spin total du système constitué par le neutron et le noyau atomique [Squires 1978, Lovesey, 1984]. Le signe et la valeur de b varient de manière irrégulière en fonction du nombre atomique Z et de la masse atomique A . La partie imaginaire représente l'absorption, elle est souvent négligeable.

Comme le neutron peut avoir deux états de spin, ($\pm 1/2$), un diffuseur de spin non nul, S , possède donc deux longueurs de diffusion différentes, b^\pm , correspondant aux états ($s \pm 1/2$) du système qui apparaissent de manière aléatoire. Il en résulte que la diffusion d'un neutron par un noyau atomique de spin $S \neq 0$ peut être caractérisée par une **longueur de diffusion cohérente**, b_{coh} , et une **longueur de diffusion incohérente**, b_{inc} telles que :

$$\begin{aligned} b_{coh} &= \langle b_i \rangle \\ b_{inc} &= \sqrt{\langle b_i^2 \rangle - \langle b_i \rangle^2} \end{aligned} \quad (1.24)$$

$\langle b_i \rangle$ est la moyenne sur toutes les orientations possibles des spins nucléaires et sur les distributions casuelles des isotopes i .

Physiquement, nous pouvons voir les diffusions cohérente et incohérente de la manière suivante : puisque la longueur de diffusion varie d'un isotope à l'autre, et dépend de l'orientation du spin nucléaire, le neutron ne voit pas un potentiel de diffusion uniforme mais un potentiel qui change d'un atome à l'autre. Seul le potentiel moyen, proportionnel à $\langle b_i \rangle$, peut donner des effets d'interférence et donc une diffusion cohérente. L'écart par rapport au potentiel moyen est distribué de manière aléatoire et ne peut donner aucun effet d'interférence. Il conduit alors à une diffusion incohérente qui est proportionnelle à la déviation quadratique moyenne ($\langle b_i^2 \rangle - \langle b_i \rangle^2$).

Les longueur de diffusion, b_{coh} et, b_{inc} sont reliées aux sections efficaces de diffusion cohérente et incohérente par les relations suivantes :

$$\begin{aligned} \sigma_{coh} &= 4\pi \langle b_i \rangle^2 \\ \sigma_{inc} &= 4\pi (\langle b_i^2 \rangle - \langle b_i \rangle^2) \end{aligned} \quad (1.25)$$

avec

$$\sigma_{totale} = \sigma_{coh} + \sigma_{inc}$$

Les longueurs de diffusion sont déterminées expérimentalement car, à l'heure actuelle, il n'est toujours pas possible de les évaluer théoriquement. Il est intéressant de remarquer que les longueurs de diffusion de deux isotopes sont, en général, très différentes. La substitution isotopique est exploitée dans certaines expériences pour modifier l'intensité relative de

diffusion des différents constituants d'un système sans en modifier significativement les propriétés chimiques.

Il est très important de remarquer les différences substantielles qui existent entre les sections de diffusion cohérente, σ_{coh} , et incohérente, σ_{inc} , de l'hydrogène (H) et du deutérium (D).

Pour l'hydrogène, dont le spin nucléaire est de 1/2, $b^+ = 1.01 \cdot 10^{-12}$ cm et $b^- = -4.74 \cdot 10^{-12}$ cm

$$\sigma_{\text{coh}} = -1,76 \text{ barn et } \sigma_{\text{inc}} = 79,91 \text{ barn} \quad (1 \text{ barn} = 10^{-24} \text{ cm}^2)$$

Par contre pour le deutérium, dont le spin nucléaire est 1, $b^+ = 0.951 \cdot 10^{-12}$ cm et

$$b^- = 0.1 \cdot 10^{-12} \text{ cm}$$

$$\sigma_{\text{coh}} = 5,60 \text{ barn et } \sigma_{\text{inc}} = 2,04 \text{ barn.}$$

La très forte diffusion incohérente de l'hydrogène est une propriété dont il est nécessaire de tenir compte lors des expériences. La diffusion par un **système hydrogéné** est dominé par la **diffusion incohérente** des atomes d'hydrogène. Ceci peut être mis à profit dans les études de dynamique visant à caractériser les mouvements individuels des atomes d'hydrogène. Par contre lors des études structurales aux petits angles la diffusion incohérente donne un bruit de fond indésirable. Pour l'éviter, il sera donc indispensable d'employer un système totalement, ou très fortement, deutéré.

1.3.2.2 Intensité diffusée et facteurs de structure

Nous allons maintenant donner les équations, qui permettent d'exprimer l'intensité diffusée par un échantillon.

Dans le cas simple d'un système constitué d'un seul type d'élément et de N diffuseurs, l'utilisation de la règle de Fermi (Squire, 1978), et du potentiel d'interaction de Fermi, conduit à l'expression suivante de la **section efficace différentielle**:

$$\left(\frac{d^2\sigma}{d\Omega d\omega} \right) = \frac{1}{N} \frac{k_i}{k_f} \frac{1}{2\pi\hbar} \sum_{jj'} \overline{b_j b_{j'}} \int_{-\infty}^{+\infty} \langle \exp(-i\mathbf{QR}_{j'}(0)) \exp(i\mathbf{QR}_j(t)) \rangle \exp(-i\omega t) dt$$

(1. 26)

où :

les crochets indiquent qu'une moyenne statistique est effectuée sur tout le système.

$\mathbf{R}_j(t)$ est l'opérateur de position de l'atome j à l'instant t

$b_j b_{j'}$ vérifie la relation :

$$\overline{b_j b_{j'}} = \langle b \rangle^2 + \delta_{jj'} (\langle b^2 \rangle - \langle b \rangle^2) \quad (1.27)$$

La section efficace différentielle de diffusion d'un système peut être décomposée en deux termes, le premier correspondant à la diffusion cohérente et l'autre à la diffusion incohérente (Squires, 1978 ; Lowesey 1984, Bée 1988 ; Als-Nielsen, 1993):

$$\left(\frac{d^2\sigma}{d\Omega d\omega} \right) = \frac{k_f}{k_i} \frac{\sigma_{coh}}{4\pi} S(\mathbf{Q}, \omega) + \frac{k_f}{k_i} \frac{\sigma_{inc}}{4\pi} S_{inc}(\mathbf{Q}, \omega) \quad (1.28)$$

où $S_{inc}(\mathbf{Q}, \omega)$ et $S(\mathbf{Q}, \omega)$ sont les *fonctions de diffusion* ou *facteurs de structure dynamique* cohérente et incohérente respectivement, définies par :

$$\begin{aligned} S(\mathbf{Q}, \omega) &= \frac{1}{2\pi\hbar N} \int \exp(-i\omega t) \sum_{jj'} \langle \exp(-i\mathbf{Q}\mathbf{R}_{j'}(0)) \exp(i\mathbf{Q}\mathbf{R}_j(t)) \rangle dt \\ S_{inc}(\mathbf{Q}, \omega) &= \frac{1}{2\pi\hbar N} \int \exp(-i\omega t) \sum_j \langle \exp(-i\mathbf{Q}\mathbf{R}_j(0)) \exp(i\mathbf{Q}\mathbf{R}_j(t)) \rangle dt \end{aligned} \quad (1.29)$$

Ces fonctions sont les transformées de Fourier en énergie des : *fonctions intermédiaires de diffusion* cohérente $I(\mathbf{Q}, t)$ et incohérente $I_{inc}(\mathbf{Q}, t)$:

$$\begin{aligned} I(\mathbf{Q}, t) &= \frac{1}{N} \sum_{jj'} \langle \exp(-i\mathbf{Q}\mathbf{R}_{j'}(0)) \exp(i\mathbf{Q}\mathbf{R}_j(t)) \rangle \\ I_{inc}(\mathbf{Q}, t) &= \frac{1}{N} \sum_j \langle \exp(-i\mathbf{Q}\mathbf{R}_j(0)) \exp(i\mathbf{Q}\mathbf{R}_j(t)) \rangle \end{aligned} \quad (1.30)$$

Comme le montrent l'expression de la (1. 29), la diffusion cohérente résulte des corrélations spatio-temporelles existant entre les atomes i et j . Par contre la diffusion incohérente provient des seules corrélations temporelles des diffuseurs.

Dans le cas le plus général où le diffuseur est une macromolécule constituée de plusieurs types d'atomes, l'équation (1. 28) se généralise aisément en effectuant au préalable la somme sur tous les atomes constituant chaque molécule (Bée, 1988)

1.3.2.3 La diffusion aux petits angles

La diffusion des neutrons aux petits angles concerne les faibles valeurs du transfert de vecteur d'onde Q . Elle est principalement utilisée pour étudier la structure d'objets de taille relativement importante, comme les macromolécules biologiques ou non. Cette diffusion peut être considérée comme élastique [Feigin L. A. and Svergun D. I., 1989,]. L'expression (1. 29), devient :

$$\frac{d\sigma}{d\Omega} = \frac{1}{4\pi} (\sigma_{inc} S_{inc}(Q) + \sigma_{coh} S(Q)) \quad (1. 31)$$

avec

$$S(Q) = I(Q, t = 0) = \frac{1}{N} \sum_{i,j} \langle \exp(i\mathbf{Q} \cdot \mathbf{r}_{ij}) \rangle \quad (1. 32)$$

où $\mathbf{r}_{ij} = \mathbf{R}_i(0) - \mathbf{R}_j(0)$.

$$\text{et} \quad S_{inc}(Q) = 1 \quad (1. 33)$$

Quand Q tend vers zéro, la fonction de diffusion $S(Q)$ tend vers N , et la fonction de diffusion incohérente est égale à 1. Suivant les valeurs du contraste et des sections efficaces totales cohérente et incohérente, le terme $\sigma_{coh} S(Q)$ peut devenir prépondérant. En fait, en diffusion aux petits angles, c'est la **diffusion cohérente** qui est observée.

La **diffusion cohérente**, intégrée sur toutes les énergies, a une distribution spatiale caractéristique de la distribution des atomes dans l'échantillon et donne des **informations sur la structure d'un composé**. Il est ainsi possible de déterminer la structure et la taille de la molécule, à basse résolution.

La **diffusion incohérente** intégrée sur toutes les énergies donne un **bruit de fond** constant en Q , qui est proportionnel à la section efficace de diffusion incohérente. Sa valeur peut être calculée théoriquement pour chaque échantillon, sachant le nombre d'hydrogènes qui sont présents.

Pour **mesurer expérimentalement** le signal de diffusion aux petits angles, on mesure tous les neutrons qui sont diffusés dans un angle solide sans tenir compte de l'énergie échangée entre le neutron et le système diffuseur (approximation élastique). On suppose que $S(Q) = S(Q,0)$ c'est-à-dire que la moyenne d'ensemble de la structure coïncide avec la moyenne sur un temps infini.

Les principes et le formalisme de la diffusion aux petits angles seront développés dans le paragraphe 1.3.4.

1.3.2.4 Diffusion incohérente inélastique des neutrons.

Pour des macromolécules biologiques naturellement fortement hydrogénées, la section efficace de diffusion de l'équation (1.28) est dominée par la *diffusion incohérente due aux atomes d'hydrogène*. Ceci provient du fait que les atomes d'hydrogène, comme nous venons de le démontrer, ont une section efficace plus de 40 fois supérieure à celle des autres atomes qui forment la molécule.

Habituellement on distingue trois régions dans un spectre de diffusion incohérente des neutrons, en fonction de l'énergie le *pic élastique*, la *diffusion quasi-élastique* et la *diffusion inélastique*. (Fig 1. 4). Le *pic élastique* représente le nombre de neutrons qui ont traversé l'échantillon sans avoir échangé d'énergie ; il constitue en général la partie la plus intense du spectre. La *diffusion quasi-élastique* est due essentiellement aux mouvements de type diffusif.

La *diffusion inélastique* provient des excitations vibrationnelles (Bée, 1988). Il faut souligner que la séparation entre la région inélastique et la région quasi-élastique est très délicate. En effet, certains mouvements diffusifs peuvent être suffisamment rapides pour s'étendre jusque dans la région inélastique. Inversement, certains mouvements de vibration sont suramortis et sont alors détectés dans la région quasi-élastique.

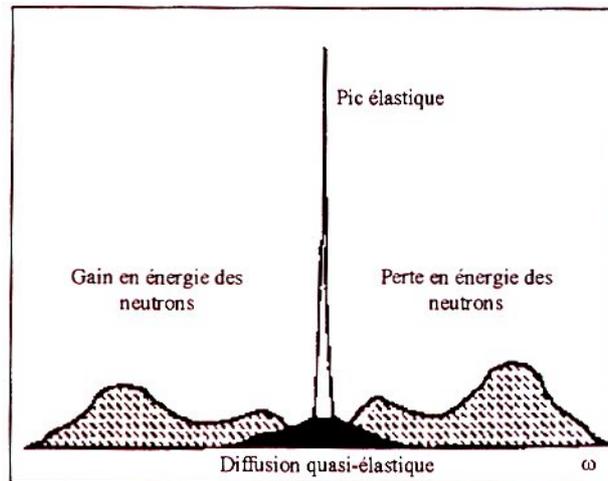


Fig 1. 4 Représentation schématique des trois types de diffusion observés en diffusion incohérente de neutrons : élastique, quasi-élastique, inélastique (Cusak, 1989)

1.3.2.4.1 Fonction d'autocorrelation.

La fonction de diffusion incohérente $S_{inc}(\mathbf{Q}, \omega)$ est la transformée spatio-temporelle de la fonction d'autocorrélation de Van Hove (Squires, 1978 ; Lowesey 1984):

$$S_{inc}(\mathbf{r}, t) = \frac{1}{2\pi\hbar} \int \exp(-i\mathbf{Q}\mathbf{r}) \exp(-i\omega t) G_{inc}(\mathbf{r}, t) d\mathbf{r} dt \quad (1.34)$$

où :

$$G_{inc}(\mathbf{r}, t) = \frac{1}{N} \sum_j \langle \delta(\mathbf{r} - \mathbf{R}_j(0) + \mathbf{R}_j(t)) \rangle \quad (1.35)$$

$G_{inc}(\mathbf{r}, t)$ est la probabilité pour qu'une particule située à l'origine au temps $t=0$, se trouve au temps t dans la position \mathbf{r} . Cette grandeur permet de donner des informations directes sur **l'évolution temporelle des positions des atomes**, c'est à-dire sur leur dynamique.

Pour $t \rightarrow \infty$ la corrélation entre $R(0)$ et $R(t)$ devient indépendante du temps. Ainsi la relation (1.35) peut être écrite comme la somme de deux composantes :

$$\begin{aligned} G_{inc}(\mathbf{r}, t) &= G_{inc}(\mathbf{r}, \infty) + G'_{inc}(\mathbf{r}, t) \\ \lim_{t \rightarrow \infty} G'_{inc}(\mathbf{r}, t) &= 0 \end{aligned} \tag{1.36}$$

ce qui permet de séparer la partie élastique de la partie inélastique.

La fonction de diffusion incohérente peut par conséquent être divisée en deux parties :

$$S_{inc}(Q, \omega) = \int G_{inc}(\mathbf{r}, \infty) \delta(\omega) \exp(-i\mathbf{Q}\mathbf{r}) d\mathbf{r} + S'_{inc}(Q, \omega) \tag{1.37}$$

La première partie, est un pic de Dirac, il correspond au **pic élastique** dans la courbe de diffusion. Il a les dimensions d'un facteur de structure et est appelé '**facteur de structure élastique incohérent**' (EISF). Il décrit l'ensemble des positions que peut atteindre l'atome pendant un temps infini, par conséquent il donne des informations directes sur la **géométrie de l'espace accessible** à ce diffuseur. L'existence d'un pic élastique traduit la présence d'un diffuseur dont les mouvements sont localisés dans l'espace. S'il n'y a pas de pic élastique, il s'agit d'un liquide (Bée, 1988).

Le second terme correspond à la **diffusion inélastique** et **quasi-élastique**, qui rend compte directement de la **dynamique** de la molécule.

Nous allons discuter ces grandeurs de façon plus détaillée au cours du chapitre V.

1.3.3 Diffusion des rayons X

Les rayons X constituent un rayonnement électromagnétique de longueur d'onde proche de l'Angström. L'énergie E associée aux photons égale à :

$$E = \hbar\omega = \frac{hc}{\lambda} \quad (1.38)$$

Les rayons X interagissent avec les charges électriques qu'ils mettent en vibration à la même fréquence que la leur. Il en résulte une diffusion dipolaire. La section efficace différentielle de diffusion par une particule peut s'écrire :

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{m^2 c^4} \left(\frac{1 + \cos^2 \theta}{2} \right) \quad (1.39)$$

où c est la vitesse de la lumière, e est la charge de l'électron, m est la masse de la particule chargée.

Puisque $d\sigma/d\Omega$ est inversement proportionnelle à m^2 , la diffusion due aux protons est négligeable. Les rayons X interagissent donc principalement avec le cortège électronique des atomes et cette interaction est caractérisée par la grandeur b_i , équivalente à la longueur de diffusion b_i des neutrons. (Williams, 1993):

$$b_e = (e^2/mc^2) f_Z(Q) \quad (1.40)$$

où $f_Z(Q)$ est le facteur de forme atomique.

Pour un atome de numéro atomique Z l'amplitude du facteur de forme est (James, 1962) :

$$f_Z(Q) = \frac{1}{Z} \int \rho_Z(\mathbf{R}) \exp(i\mathbf{Q}\cdot\mathbf{R}) d\mathbf{R} \quad (1.41)$$

où $\rho_Z(\mathbf{R})$ est la densité électronique de l'atome. Pour $Q \Rightarrow 0$ le facteur de structure atomique tend vers le numéro atomique Z . L'expression numérique de la longueur de diffusion est (en cm) (Cotton, 1991) :

$$b_i = 0.282 \cdot 10^{-12} Z \quad (1.42)$$

Cette longueur de diffusion est proportionnelle à Z . Il en résulte que l'oxygène ou l'azote ou le carbone ont des longueurs de diffusion semblables, six à huit fois plus élevées que celle de l'hydrogène, qui contribuera peu à la diffusion

Les atomes ont donc des comportements différents suivant qu'ils interagissent avec des neutrons ou des rayons X. Par exemple, les neutrons sont très sensibles à la présence d'hydrogène et de deutérium au sein des diffuseurs, alors que les rayons X ne le sont que très peu. D'autre part, les rayons X ne donnent pas de diffusion incohérente. Toutefois, en ce qui concerne la diffusion aux petits angles le traitement théorique, exposé au chapitre suivante, est le même pour les neutrons et les rayons X.

1.3.4 Formalisme de la diffusion aux petits angles

1.3.4.1. Diffusion par une solution idéale de macromolécules identiques

L'étude des changements structuraux de la NCS au cours de la dénaturation a été suivie essentiellement par diffusion aux petits angles. Comme une molécule peut adopter toutes les orientations possibles lorsqu'elle est en solution, il n'est pas possible d'obtenir des informations au niveau atomique (Jacrot B. 1976). Par contre, il est possible de déterminer sa taille et sa forme et, dans certains cas, sa structure interne.

Considérons un ensemble N de macromolécules en solution. Comme elles peuvent y prendre toutes les orientations possibles, la diffusion est isotrope. Si le solvant est constitué de molécules de taille bien plus petite, il peut être considéré comme un milieu continu. Dans l'hypothèse d'une solution monodisperse idéale, c'est-à-dire pour laquelle les particules sont *identiques et sans interactions*, le soluté produit, par rapport au solvant, un excès de diffusion cohérente :

$$I_{\text{coh}}(Q) = \frac{cM}{N_A} K^2 P(Q) \quad (1.43)$$

où c est la concentration des macromolécules, M leur masse molaire, N_A est le nombre d'Avogadro.

K est le contraste moyen de la molécule par rapport au solvant. Sa valeur est :

$$K = \langle \rho_p \rangle v - \langle \rho_s \rangle v_P \quad (1.44)$$

où $\langle \rho_p \rangle$ et $\langle \rho_s \rangle$ sont les densités de longueur de diffusion (cm^{-2}) de la molécule et du solvant, respectivement. Nous définissons $\langle \rho_p \rangle = \Sigma b_i / V_p$, avec b_i la longueur de diffusion et V_p le volume de la protéine. v est le volume spécifique de la molécule et v_p son volume spécifique partiel ($\text{cm}^3 \text{g}^{-1}$).

Dans le cas de la diffusion des neutrons par des protéines hydrogénées, le contraste est maximal quand le solvant est constitué à 100% d'eau lourde.

Enfin, $P(Q)$, le facteur de forme de la molécule en solution, s'écrit:

$$P(Q) = \langle |A(Q)|^2 \rangle = \left\langle \left| \frac{1}{K} \int (\rho(r) - \rho_s) \exp(i\mathbf{Q} \cdot \mathbf{r}) d\mathbf{r} \right|^2 \right\rangle \quad (1.45)$$

$\langle \dots \rangle$ désigne la moyenne sur toutes les orientations de la molécule.

Approximation de Guinier et rayon de giration

Aux faibles valeurs de Q le facteur de forme se simplifie par l'approximation de Guinier :

$$P(Q) \cong \exp[-(QR_g)^2/3] \quad (1.46)$$

où R_g est le rayon de giration de la macromolécule. Le domaine de validité de cette approximation dépend de la forme de cette dernière. Pour un objet compact il s'étend jusqu'à $Q \leq R_g^{-1}$. Pour un objet étendu (comme une protéine dénaturée), le domaine de validité de l'approximation de Guinier est tellement réduit qu'il est alors nécessaire de faire appel à une autre approximation (Debye). Ceci sera détaillé dans la section suivante.

Le rayon de giration est défini par l'expression :

$$R_g^2 = \frac{1}{KV_p} \int \langle \rho(\mathbf{r}) - \rho_s \rangle r^2 d\mathbf{r} \quad (1.47)$$

Le carré du rayon de giration est donc la moyenne des carrés des distances r des atomes de l'objet par rapport à son centre de gravité, pondérés par le contraste. Ceci implique que les valeurs des rayons de giration ne seront pas exactement les mêmes si le rayonnement est constitué de neutrons ou de rayons X, dans la mesure où le contraste est différent dans les deux cas.

La valeur de R_g donne une estimation de la taille du diffuseur. Si ce dernier n'est pas homogène, son rayon de giration dépend du contraste.

L'intensité à l'origine

Le second paramètre intéressant, des courbes de diffusion, qui peut être analysé avec profit est la valeur $I(0)$. Cependant puisque cette valeur ne peut pas être mesurée directement, il est nécessaire de faire une extrapolation à $Q=0$. L'intérêt de $I(0)$ résulte de sa liaison avec la masse moléculaire et le volume moléculaire (Jacrot 1976). Dans le cas d'une solution diluée, l'expression s'écrit (Zaccai & Jacrot 1983)

$$I(0) = \frac{Mc}{N_A} \left(\sum_i b_i - \rho_s v_p \right)^2 \quad (1.48)$$

c désigne la concentration de la protéine, v_p le volume spécifique de la protéine, M la masse molaire, N_A le nombre d'Avogadro.

L'équation (1.48) suggère une méthode pour évaluer la valeur v_p à partir des données de diffusion. La valeur de $I(0)$ doit aussi être contrôlée pour suivre d'éventuels phénomènes d'agrégation.

1.3.4.2. Diffusion par un mélange de macromolécules identiques en solution diluée.

Tout ce qui précède concerne l'intensité diffusée par une solution idéale de particules identiques. Les expressions deviennent plus compliquées lors de la dénaturation d'une protéine, puisqu'il y a alors en présence au moins deux populations : la protéine native et la protéine complètement dépliée, et éventuellement des protéines dans des états intermédiaires. L'intensité diffusée s'écrit alors sous la forme :

$$I(Q) = \sum f_i I_i(Q), \quad (1.49)$$

Où le f_i sont les fractions des différentes populations, et les $I_i(Q)$ les intensités de diffusion correspondantes. Les quantités mises en jeu ci-dessus sont dépendantes du paramètre, qui varie au cours de la dénaturation, par exemple la température T . L'équation (1.49) peut être réécrite sous la forme :

$$I(Q,T) = \sum f_i(T) I_i(Q,T), \quad (1.50)$$

La détermination des courbes expérimentales $I(Q,T)$ mesurées pour différentes valeurs du paramètre T permet d'obtenir des informations sur la structure des différents états en présence, à condition de séparer l'effet dû à T dans les courbes de diffusion $I_i(Q,T)$ de chacune des populations, c'est-à-dire à condition de mettre $I_i(Q,T)$ sous la forme :

$$I_i(Q,T) = K_i(T, v_{pi}) P_i(Q), \quad (1.51)$$

où $P_i(Q)$ est le facteur de forme et $K(T, v_{pi})$ le terme de contraste.

Si on suppose que les volumes spécifiques des différentes espèces ont des valeurs suffisamment proches pour que le terme de contraste puisse être considéré comme dépendant seulement de T , l'équation (1.49) devient :

$$I(Q,T) / I(0,T) = \sum f_i(T) P_i(Q), \quad (1.52)$$

Le carré du rayon de giration mesuré à la température T est donc :

$$R_g^2(T) = \sum_i f_i R_i^2(T), \quad (1.53)$$

où R_i est le rayon de giration de l'espèce i .

1.3.4.3 Facteurs de forme des chaînes polymériques

Dans ce paragraphe nous résumons quelques résultats de la théorie des polymères concernant en particulier le facteur de forme $P(Q)$. Cela nous sera très utile par la suite, car le spectre de diffusion d'une solution de protéines totalement dépliées est tout à fait semblable à celui d'une solution diluée de polymères.

Nous rappelons quelques *concepts de base sur les polymères* dans la section encadrée.

Un homopolymère est une molécule linéaire formée d'une série d'éléments chimiques identiques, les monomères. Un polymère en solution peut se comporter de différentes manières selon le type de solvant dans lequel il se trouve. (de Gennes 1979)

- si les interactions polymère-polymère sont équivalentes aux interactions polymère-solvant, le polymère se trouve alors en solvant θ . Il peut adopter toutes les conformations statistiquement possibles et la distribution des monomères est gaussienne. Il se comporte alors comme une *chaîne gaussienne* ou *idéale* ;
- dans le cas où les interactions polymère-solvant sont privilégiées, le polymère est totalement solvaté. Le polymère se trouve alors en bon solvant et ne peut plus adopter toutes les conformations qu'il aurait en solvant θ . Il se comporte alors comme une *chaîne à volume exclu* ;
- dans le cas où les interactions intramoléculaires sont privilégiées, le polymère se trouve en mauvais solvant. Il adopte alors une *conformation effondrée plus ou moins compacte*.

Une protéine est un hétéropolymère dont les monomères sont des résidus d'acides aminés. En bon solvant les différences entre les divers résidus sont marginales et une protéine se comportera comme une *chaîne à volume exclu*. Par contre au voisinage du point θ et surtout en mauvais solvant, des interactions préférentielles entre certains résidus apparaîtront et la protéine ne se comportera plus comme un homopolymère.

Loi de Debye

Le modèle gaussien est mathématiquement le modèle le plus simple. Une expression analytique, qui décrit le profil de diffusion pour une chaîne gaussienne a été déterminée par Debye. (Debye 1944) :

Cette loi permet de déterminer le rayon de giration dans un domaine de vecteur de diffusion qui satisfait la relation $QR_g < 3$:

$$P_D(Q) = \frac{2}{x^2} [x - 1 + \exp(-x)] \quad (1.54)$$

où $x = (QR_g)^2$.

Cette loi peut également être appliquée pour déterminer le rayon de giration, pour des configurations à volume exclu (Rawiso 1987), puisque dans la zone de Q explorée le comportement de la chaîne étendue par rapport à son solvant ne modifie pas l'allure de la courbe de diffusion.

Chaîne à longueur de persistance

Lorsque $Q \geq 3(R_g)^{-1}$ la fonction de Debye (1.54) ne donne pas une représentation correcte du spectre de diffusion d'une chaîne polymérique réelle. Les modèles qui permettent de décrire la conformation moyenne d'une telle chaîne se classent naturellement en fonction de la résolution spatiale, Q^{-1} . A haute résolution, où les détails de la conformation et de la structure se révèlent, il est nécessaire de tenir compte des interactions existant entre différentes parties de la chaîne. Les interactions à longue distance sont dues aux effets de volume exclu. En revanche, les interactions à courte distance résultent essentiellement d'un effet stérique, qui est la restriction à la rotation autour des liaisons de la chaîne principale.

La chaîne présente une rigidité locale qui est représentée par la longueur statistique b . La longueur statistique est la longueur minimale d'un segment dont l'une des extrémités peut prendre toutes les orientations possibles par rapport à l'autre. Elle permet de distinguer les polymères rigides des polymères flexibles à l'aide du rapport b/L , où L est la longueur de contour du filament. Une chaîne flexible correspond à $b/L \ll 1$ et une chaîne rigide à $b/L \gg 1$.

Si des Cloiseaux (1973) a donné une expression du facteur de forme d'une chaîne semi-flexible infiniment longue ($L/b \rightarrow \infty$), l'expression de Sharp et Bloomfield (1968) :

$$P_{SB}(x) = P_D(x) + \frac{b}{L} \left[\frac{4}{15} + \frac{7}{15x} - \left(\frac{11}{15} + \frac{7}{15x} \right) \exp(-x) \right] \quad (1.55)$$

où $P_D(x)$ est la fonction de Debye, $x = Q^2 Lb/6$, est plus appropriée pour décrire une chaîne de longueur finie, comme une protéine dénaturée. Cependant ce résultat n'est valable que pour une **chaîne idéale** telle que $L/b > 10$ et $Qb < 3,1$.

Le rayon de giration de la chaîne à longueur de persistance a été calculé par Benoit-Doty (1953)

$$R_g^2 = b^2 \left[\frac{y}{6} - \frac{1}{4} + \frac{1}{4y} - \frac{1}{8y^2} (1 - \exp(-2y)) \right] \quad (1.56)$$

$$y = \frac{L}{b}$$

Lorsque $Qb > 3,1$ les difficultés théoriques deviennent trop importantes pour obtenir directement le facteur de forme d'une chaîne semi-flexible de longueur finie. En utilisant la méthode de Monte Carlo, Pedersen et Schurtenberger (1996) ont d'abord simulé les spectres de diffusion de telles chaînes, avec et sans volume exclu. Ils ont ensuite représenté les résultats des simulations par diverses expressions analytiques qui peuvent être utilisées pour interpréter les spectres expérimentaux. Les fonctions de diffusion sont approchées, aux petits valeurs de Q , avec une chaîne gaussienne ou volume exclu, et se superposent avec une expression pour un *cylindre* pour les valeurs de Q plus grandes. Avec ce modèle il est possible d'étudier la zone de Q intermédiaire ($1/Rg < Q < 1/a$ avec a la taille caractéristique d'un centre diffuseur), qui nous permet de mieux caractériser la conformation de la chaîne.

Les expressions correspondant à la 3^{ème} approximation de P&S sont les suivants :

- dans le cas d'**une chaîne idéale** l'expression (1.55) est utilisée pour les valeurs de Q telles que $Qb \leq 3,1$. Pour les valeurs supérieures, l'approximation suivante est utilisée :

$$P_{PS}(QL, Qb) = \frac{a_1}{(Qb)^{p_1}} + \frac{a_2}{(Qb)^{p_2}} + \frac{\pi}{QL} \quad (1.57)$$

où $p_1 = 4,95$ et $p_2 = 5,13$ sont des constantes ajustées de telle sorte que l'expression décrive correctement le résultat des simulations. Par contre, les coefficients a_1 et a_2 doivent être calculés en imposant la continuité du facteur de forme,

$$P(QL, Qb) = P_{SB}(QL, Qb) + P_{PS}(QL, Qb), \quad (1. 58)$$

et de sa dérivée par rapport à Q pour $Q = 3,1b^{-1}$;

- dans le cas d'une **chaîne à volume exclu**, l'expression (1. 57) demeure valable lorsque $Qb > 3,1$. Par contre, aux plus faibles valeurs de Q , l'expression de Sharp et Bloomfield est modifiée de la manière suivante :

$$P_{SB,VE}(x) = P_{VE}(x) + a_4 \left(\frac{b}{L} \right)^{1+p_3} \left[\frac{4}{15} + \frac{7}{15x} - \left(\frac{11}{15} + \frac{7}{15x} \right) \exp(-x) \right] \quad (1. 59)$$

Maintenant, l'expression du rayon de giration est

$$R_g^2 = \frac{Lb}{6} \left[1 + \left(\frac{L}{3.12b} \right)^2 + \left(\frac{L}{8.67b} \right)^3 \right]^{(2\nu-1)}, \quad (1. 60)$$

où $\nu = 0,588$ est l'exposant de volume exclu. Le premier terme de l'expression (1. 59) est

$$P_{VE}(x) = w(x) P_D(x) + [1 - w(x)] (C_1 x^{-1/\nu} + C_2 x^{-2/\nu} + C_3 x^{-3/\nu}) \quad (1. 61)$$

où

$$w(x) = \frac{1}{2} \left[1 - \tanh \frac{x - C_4}{C_5} \right] \quad (1. 62)$$

est une fonction de raccordement empirique. Dans les expressions précédentes les constantes numériques valent $a_4 = 3,06$, $p_3 = 0,44$, $C_1 = 1,220$, $C_2 = 0,4288$, $C_3 = -1,651$, $C_4 = 1,523$ et $C_5 = 0,1477$.

Enfin, que la chaîne présente ou non des interactions de volume exclu, il est nécessaire de tenir compte de ses dimensions transverses. En première approximation, ceci peut être fait en multipliant les expressions complètes des facteurs de forme précédents par celui de la section de la chaîne [Rawiso et al.]. Si, en moyenne, la chaîne peut être considérée comme cylindrique le facteur de forme de sa section est donnée par la fonction d'Airy :

$$P_S(Q) = [2J_1(QR_c)/(QR_c)]^2 \quad (1.63)$$

où $J_1(y)$ est la fonction de Bessel du premier ordre et R_c le rayon transverse. Pour des valeurs de Q suffisamment faibles, ce facteur de forme peut être décrit par une approximation de Guinier

$$P_S(Q) \cong \exp[-(QR_{g,c})^2/2] \quad (1.64)$$

où $R_{g,c}$ est le rayon de giration de la section transverse de la chaîne (Rawiso et al.). Lorsque que cette dernière est cylindrique $(R_{g,c})^2 = (R_c)^2/2$.

1.3.4.4 Représentation de Kratky

Lors de l'étude des polymères, il est courant de présenter les courbes de diffusion dans la représentation de Kratky. Elle consiste à tracer la quantité $Q^2I(Q)$ en fonction de Q . Elle permet d'obtenir immédiatement des informations qualitatives sur la conformation de la chaîne en examinant la dépendance de l'intensité de diffusion en fonction du vecteur de diffusion.

En effet les courbes sont très différentes selon que le diffuseur est un globule compact ou une chaîne dépliée.

- La courbe de diffusion d'un globule compact suit à grand Q la loi de Porod (au moins approximativement) $I(Q) \propto Q^{-4}$. En représentation de Kratky la courbe présente une décroissance hyperbolique en Q^{-2} pour des valeurs de Q suffisamment grandes (Porod 1951). La représentation de Kratky présente un pic, dont la position dépend de la valeur de R_g .

- Pour une chaîne dépliée, la dépendance en Q de l'intensité diffusée, en représentation de Kratky, est beaucoup moins rapide que pour un globule compact. Il n'apparaît plus de pic bien marqué. Toutefois il est difficile de donner des détails sur la nature de la chaîne simplement en observant l'allure de la courbe dans la représentation de Kratky, car elle dépend de plusieurs paramètres :
 - de la valeur du rayon de giration, R_g ;
 - de la présence ou l'absence d'interactions de volume exclu ;
 - des valeurs de la longueur de contour, L , de l'élément statistique, b , et du rayon de giration moyen, $R_{g,c}$, de la section de la chaîne. La valeur de cette dernière quantité dépend du contraste locale entre la protéine et le solvant (R_{wiso}).

Selon la conformation de la molécule et les valeurs de Q , le tableau ci-dessous récapitule différents comportements caractéristiques de l'intensité diffusée par une molécule.

Chaîne idéale	$3(R_g)^{-1} < Q < 2b^{-1}$	$Q^2 I(Q) = \text{cste}$
Chaîne à volume exclu	$3(R_g)^{-1} < Q < 2b^{-1}$	$Q^2 I(Q) = Q^{-0,3}$
Chaîne semi-flexible	$Q > 2b^{-1} ; L \gg b$	$Q^2 I(Q) \propto QL^{-1} \exp[-(QR_{g,c})^2/2]$
Molécule globulaire compacte	$Q \gg (R_g)^{-1}$	$Q^2 I(Q) \propto Q^{-2}$

1.3.4.5. Effet des interactions

Les résultats précédents concernent des solutions idéales, c'est-à-dire pour lesquelles les particules peuvent être considérées sans aucune interaction entre elles.

En fait, une solution monodisperse réelle de particules sphériques de protéines doit être décrite comme la convolution entre la distribution des particules en solution $d(\mathbf{r})$ et la distribution du contraste $K(\mathbf{r})$ (vachette 1996). L'intensité diffusée par une telle solution est alors le produit de l'intensité diffusée par une particule en solution idéale (*facteur de forme*) par le *facteur de structure de la solution*, qui tient compte de l'interaction entre les particules.

$$I(Q, c) = \frac{cM}{N_A} K^2 P(Q, 0) * S(Q, c)$$

(1. 65)

L'expression générale du facteur de structure est :

$$S(Q, c) = 1 + \frac{cN_A}{M} \left\langle \int [g(\mathbf{r}) - 1] \exp(i\mathbf{Q} \cdot \mathbf{r}) d\mathbf{r} \right\rangle \quad (1. 66)$$

où $\mathbf{r} = (\mathbf{r}_{0,1} - \mathbf{r}_{0,2})$ est le vecteur distance entre deux points particuliers des molécules 1 et 2. Pour des molécules compactes, ces points sont, en général, choisis comme étant leurs centres de masse. $g(\mathbf{r})$ est la fonction de corrélation de paire correspondante.

Lorsque $Q = 0$, $P(Q, 0) = 1$ et

$$S(c, 0) = 1 + \frac{cN_A}{M} \left\langle \int [g(\mathbf{r}) - 1] d\mathbf{r} \right\rangle = \frac{cRT}{M} K_T \quad (1. 67)$$

où R est la constante des gaz parfaits et $K_T = (c \partial \Pi / \partial c |_{T,P})^{-1}$ est la compressibilité osmotique de la solution dont la pression osmotique est :

$$\Pi = (RT/M) c (1 + A_2 M c + A_3 M c^2 + \dots) \quad (1. 68)$$

où A_2, A_3, \dots sont le second, le troisième, ... coefficients du viriel. Lorsque les interaction entre diffuseurs sont répulsives les coefficients du viriel sont positifs. Ils sont négatifs dans le cas contraire.

Les expressions (1. 67) (1. 68) montrent alors que l'intensité cohérente diffusée vers l'avant en présence d'interactions, peut s'écrire de la manière suivante :

$$I_{\text{coh}}(0, c) = \frac{cM}{N_A} K^2 (1 + 2A_2 M c + 3A_3 M c^2 + \dots)^{-1} \quad (1. 69)$$

A $Q \neq 0$ il est possible de mettre l'intensité diffusée par une solution de macromolécules sous la forme approchée, vraie au premier ordre en Q^2

$$I_{\text{coh}}(Q, c) \cong \frac{cM}{N_A} K^2 (1 + 2A_2 M c + 3A_3 M c^2 + \dots)^{-1} P(Q, c)$$

$$(1.70)$$

où

$$P(Q,c) \cong 1 - [QR_g(c)]^2/3 \quad (1.71)$$

est un facteur de forme apparent. $R_g(c)$ est le rayon de giration apparent que l'on mesure expérimentalement. Son carré varie de la manière suivante avec la concentration :

$$[R_g(c)]^2 = (R_g(0))^2 (1 + 2B_2Mc + 3B_3Mc^2 + \dots)^{-1}. \quad (1.72)$$

où les coefficients B_2, B_3, \dots sont en général différents des coefficient du viriel. Ces résultats ne sont valables que pour de faibles valeurs de Q , telles que $QR_g(c) \ll 1$.

Correction du facteur de forme par la concentration

En résumé les expressions ci-dessus montrent que, si l'on veut obtenir le facteur de forme $P(Q)$ d'une protéine en solution, à partir des mesures de l'intensité diffusée $I_{\text{coh}}(Q,c)$ par la solution réelle (non idéale), il est nécessaire de tenir compte du facteur de structure $S(Q,c)$ de la solution.

Afin de relier $I_{\text{coh}}(Q,c)$ et $P(Q)$, Zimm (1948) a proposé l'approximation suivante :

$$\frac{cMK^2}{N_A I_{\text{coh}}(Q,c)} \cong \frac{1}{P(Q)} + 2A_2Mc + \dots \quad (1.73)$$

pour représenter le spectre de diffusion d'une solution de polymères. Au premier ordre en c , cette approximation semble justifiée si le polymère peut être décrit comme une chaîne aléatoire de dimension transverse nulle. Benoît et Benmouna (1984) ont tenté de démontrer sa validité à tout les ordres en c , mais leur démarche n'est pas convaincante.

L'approximation (1.73) implique que $B_2 = A_2$. Ceci est assez peu vraisemblable pour une chaîne réelle dont les dimensions radiales ne sont pas nulles. Nous avons donc modifié l'approximation de Zimm de la manière suivante :

$$\frac{cMK^2}{N_A I_{\text{coh}}(Q, c)} \cong \frac{1 + 2A_2Mc + \dots}{1 + 2B_2Mc + \dots} \left[\frac{1}{P(Q)} + 2B_2Mc + \dots \right] \quad (1.74)$$

afin que les comportements de l'intensité diffusée vers l'avant et du rayon de giration soient décrits par les expressions (1.69) et (1.72), respectivement.

1.3.5 Diffusion quasi élastique de lumière

Les expériences de diffusion de lumière en mode dynamique homodyne permettent de mesurer la fonction d'autocorrelation de l'intensité diffusée :

$$C(Q, t) = \langle I(Q, 0) I(Q, t) \rangle \quad (1.75)$$

$I(Q, t)$ est l'intensité diffusée à l'instant t avec un vecteur d'onde de transfert $Q = (4\pi/\lambda) \sin\theta/2$ où λ la longueur d'onde de la lumière dans le milieu et θ l'angle de diffusion. $\langle \dots \rangle$ désigne la moyenne temporelle. Cette opération est réalisée par un corrélateur qui possède 256 canaux de largeur Δt_i dont on peut choisir les valeurs de différentes façons.

Pour une solution diluée de macromolécules identiques, la fonction d'autocorrélation de l'intensité diffusée a la forme suivante :

$$C(Q, t) = \langle [I(Q, 0)]^2 \rangle (1 + B \exp(-D Q^2 t / \tau)) \quad (1.76)$$

où B un facteur expérimental compris entre 0 et 1 et D est le coefficient de diffusion de translation des molécules. Il est relié au rayon hydrodynamique R_H de la molécule par la loi de Stokes-Einstein

$$D = \frac{k_B T}{6 \pi \eta R_H} \quad (1.77)$$

où η la viscosité de la solution.

